# SNSF DATA MANAGEMENT PLAN (DMP)



#### How to fill it and what information to include?

Floriane Muller, Anouk Santos

November 2021



### Plan

- General principles about Research Data and SNSF requirements
- What is the **Data Management Plan** (DMP)?
- The **essentials** in a nutshell
- Our services to create your DMP

# **UNIGE's Policy**

- ✓ Approved by the Rectorat after consultation with the faculties and deans (June 2018)
- ✓ 1<sup>st</sup> university in Switzerland with a policy on the management of research data
- Establishment of responsibilities:
  - of the University
  - of researchers

https://www.unige.ch/researchdata/en/services/all/policy/

### What are Research Data?

Factual records (numerical scores, textual records, images and sounds) used as primary sources for scientific research, and that are commonly accepted in the scientific community as necessary to validate research findings



# One concept encompassing many things



### **Classification of Research Data**

### observational















### Sources → Data → Publications

### **Publications**

Thesis, books, articles, reports

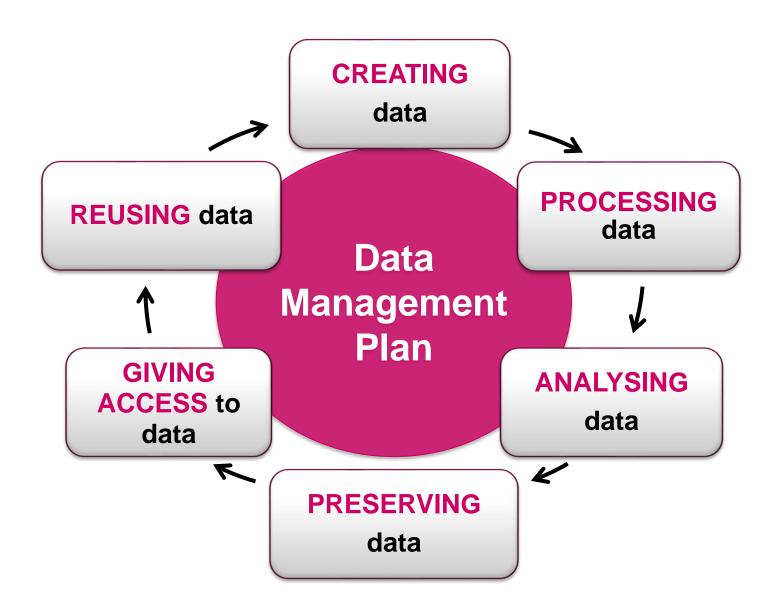
#### **Research Data**

Bibliographies, transcriptions, analyses, models, simulations.

### **Primary Sources**

Archival documents, letters, musical scores, manuscripts, sequencing data, surveys...

### **Data Life Cycle**



# **Data Management Plan**



- Saves time
- ✓ Increases the **impact** of your research
- Participates in open science
- Facilitates new discoveries



# **Requirements from SNSF**



#### 2. Application Data

#	No./Title		Status
2.1	Basic data I		In preparation
2.2	Basic data II		In preparation
2.3	Use-inspired project		In preparation
2.4	Re-submission		
2.5	Continuation of	. D	onnées relatives à la requête
2.6	Link to other SNSF projects	#	No/titre
2.7	Further requested and available funds (not from	2.1	Données de base I
2.8	University or research institution	2.2	Données de base II
2.9	Requested funding	2.3	Projet orienté vers l'application
2.10	Data management plan (DMP)	2.4	Resoumission
2.11	Research requiring authorisation or notification	2.5	Continuation
2.12	Exclusion of external reviewers	2.6	Relation avec d'autres projets FNS
2.13	General remarks on the project	2.7	Autres moyens demandés et disponibles (qui ne sont pas
		2.8	Haute école
		2.9	Besoin financier
		2.10	Plan de gestion des données (DMP)
		2.11	Recherche exigeant des autorisations ou des annonces

# The DMP on MySNF

I do not submit a DMD for the following reason:
☐ I do not submit a DMP for the following reason:
1. Data collection and documentation
☐ 1.1 What data will you collect, observe, generate or reuse?
☐ 1.2 How will the data be collected, observed or generated?
1.3 What documentation and metadata will you provide with the data?
2. Ethics, legal and security issues
2.1 How will ethical issues be addressed and handled?
2.2 How will data access and security be managed?
2.3 How will you handle copyright and Intellectual Property Rights issues?
3. Data storage and preservation
□ 3.1 How will your data be stored and backed-up during the research?
□ 3.2 What is your data preservation plan?
4. Data sharing and reuse
☐ 4.1 How and where will the data be shared?
■ 4.2 Are there any necessary limitations to protect sensitive data?
□ 4.3 All digital repositories I will choose are conform to the FAIR Data Principles.
4.4 I will choose digital repositories maintained by a non-profit organisation.

### 1. Data Collection and Documentation

1.1 What data will you collect, observe, generate or reuse?

#### List of data:

- *→ Type*
- → Format
- → Volume





Existing data to be reused in the project consists of **transcribed texts**, **facsimile images**, and **audio files** found in the third-party database Australian Drama Archive

#### https://www.austlit.edu.au/australiandramaarchive

The project itself will generate the following types of data:

- **Text transcriptions** from extracts of earlier plays which do not already exist in the Australian Drama Archive. These will be produced in the **Open XML format .docx** during the in-project period.
- Images. Where relevant the project will produce **photographic images** of some of the features of the original playscripts or ephemera identified during archival trips. An **estimated 500 images** will be taken in high resolution (2000 X 3000 pixels) as **.jpg files**.
- **Tables and figures** in **.xlsx format**, outlining different characteristics of monologues found in those earlier plays.

It is expected that the total volume of data generated **will not exceed 20 GB** but this will be further defined once archival trips are completed at the end of the first year of the project.

### 1. Data Collection and Documentation

1.2 How will the data be collected, observed or generated?

#### **Data processing:**

- → Standards
- → Methods
- → Files naming and versioning



#### v

#### 1.2 How will the data be collected, observed or generated?

All samples on which data are collected will be prepared according to **published standard protocols in the field**.

Quality of analytical data will be guaranteed through **calibration of devices** and comparison with internal standards. Appropriate experimental design, data recording and data validation (controls, randomization/ blinding, sampling/replicates, experimental versus hypothesis driven-protocol) will be used, ensuring internal validity. Methods and materials will be recorded using an **Electronic Lab Notebook (ELN).** All experimental data will be automatically imported into a **Laboratory Information Management System (LIMS)** from the measurement device.

Files will be named according to a pre-agreed **convention** including a short, unique identifier for the project, a summary of the content, the date (YYYYMMDD format) and the document version. Ex.: **YYYYMMDD\_ProjectIDX\_ContentSummary\_V01** 

Peer review will be employed as an additional step of quality control in the form of **regular** supervision and lab meetings. (eg from DMP Canva Generator, VitalIT)

### 1. Data Collection and Documentation

1.3 What documentation and metadata will you provide with the data?

- → Information required for future users
- → Annotations



#### 1.3 What documentation and metadata will you provide with the data?

A README file will be saved in the root folder of the project. It will contain the following information:

- The names of the Principal investigator and co-investigators
- A description of methodology and tools used for data collection
- The dates of data collection
- A description of the directory hierarchy, of the type of data it contains and the filenaming convention
- A complete list of any headings/codes/abbreviations and conventions used in the files
- The sharing and accessing conditions (license) during and after the end of the project

### **Metadata Sources**

- Use existing metadata standards whenever possible
- Document your metadata scheme as soon as possible
- Examples: author, title, date, brand and model of the experimental apparatus...

fome Digital curation this section A curated, informative and educational resource on data and metadata standards, inter-related to databases and data policies. HOW CAN WE HELP? We guide consumers to discover, select and use these resources with confidence, and producers to make their resource more discoverable, more widely adopted and cited. Societies, unions and community alliances Journal Publishers Librarians & Trainers

http://www.dcc.ac.uk/resources/metadata-standards https://fairsharing.org/

How to write a good README.txt

https://data.research.cornell.edu/content/readme

# 2. Ethics, Legal and Security Issues

2.1 How will ethical issues be addressed and handled?

- → Data protection
- → Confidentiality agreement
- → Personal and sensitive data management

2.1 How will ethical issues be addressed and handled?

All interviewees and focus group participants will sign a **Consent form** agreed to by **UNIGE ethics committee**.

We have guaranteed anonymity to our interviewees and focus group participants. Therefore we will not be depositing .wav files as this would compromise that guarantee.

However, **anonymised transcripts** of the interviews and focus groups will be deposited. Unless such anonymisation is impossible or associated with a disproportional effort, in which case part or all of the data will not be shared publicly. We will make sure consent forms make provision for future sharing of data.

(source of this eg., adapted)

### Which Ethics Committee?

#### If your research concerns:

- · Living or deceased persons;
- embryos and fœtuses;
- Biological material;
- Health-related personal data.

Then it falls under the HRA\*, and you should contact

Commission cantonale d'éthique et de recherche (CCER)

If your research involves human participants but does not fall within the scope of the HRA,

you should contact

UNIGE's Committee for Ethical Research (CUREG 2.0)

<sup>\*</sup> Federal Act on Research involving Human Beings

### **UNIGE's Guidelines**

#### ✓ Code of ethics and professional conduct

"Any individual engaged in research must ensure that integrity is maintained in his or her quest for knowledge, in the interpretation and application of the results and in the structuring of his or her research." (article 1, letter c)

#### ✓ Research integrity directive

The head of the research project must ensure that the essential data are kept securely for at least five years after the end of the project. (point 2.6)

- ✓ Swiss Federal Act on Research involving Human Beings (HRA)
- ✓ Commission cantonale d'éthique de recherche (CCER)
- ✓ Commission universitaire d'éthique (CUREG 2.0)

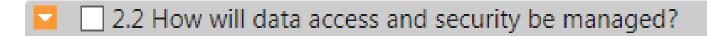
# 2. Ethics, Legal and Security Issues

2.2 How will data access and security be managed?

#### **Risks management:**

- → Data security
  - → Digital
  - → Material
- → Data access rights and permissions





Our data is stored on the academic NAS managed by the UNIGE IT department (DiSTIC). Access to the data is limited to rights holders (central authentication). The head of the laboratory that owns this disk space manages access himself, with the possibility of registering additional users.

# 2. Ethics, Legal and Security Issues

2.3 How will you handle copyright and Intellectual Property Rights issues?

#### Legal aspects:

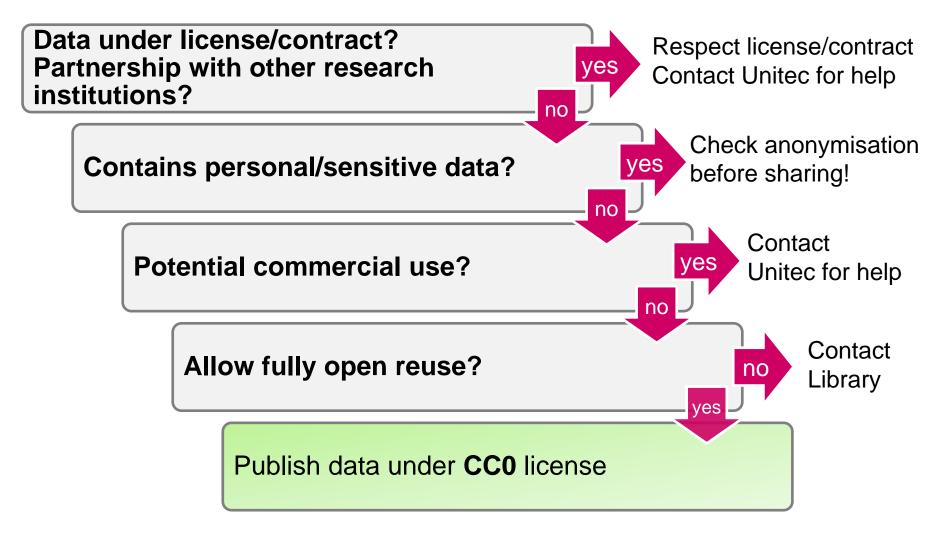
- → Who is the owner of the data?
- → Which licence to apply?
- → What restrictions apply to the reuse of third-party data?



2.3 How will you handle copyright and Intellectual Property Rights issues?

Research data generated by UNIGE collaborators in the performance of their duties is the property of the institution. As the data is not subjected to a contract and will not be patented, it will be released as open data under Creative Commons CCO license.

### **Choose your License**



See also: <a href="https://creativecommons.org/choose/">https://creativecommons.org/choose/</a>
<a href="https://creativecommons.org/choose/">https://creativecommons.org/choose/</a>
<a href="https://creativecommons.org/choose/">https://creativecommons.org/choose/</a>

# 3. Data Storage and Preservation

3.1 How will your data be stored and backed-up during the research?

#### **During the project:**

- → Storage capacity
- → Storage facility
- → Back-up procedures

3.1 How will your data be stored and backed-up during the research?

During the project, data will be stored on the **PI's computer hard drive** and backed up weekly:

- On an external hard drive kept in a locked filing cabinet in the PI's office –
   the PI will perform the backup
- On the academic NAS managed by the University of Geneva's IT department. This academic NAS follows common protocols and best practices to ensure maximum security, integrity and availability. It extends over two distinct physical locations (UniDufour and Campus Biotech) and automatically performs a snapshot of files every 4 hours, with a retention of copies of 6 weeks.

**Consent forms** and all identifying information will be kept in a locked filing cabinet and not stored with electronic files.



# **Good Practices for Data Storage**

→ Store 3 copies of your data



- 1. The original
- 2. A copy kept on a local external device
- 3. A copy kept on an external device at a different location

UNIGE's NAS servers, a shared storage space, covers points 2. and 3.

# 3. Data Storage and Preservation

3.2 What is your data preservation plan?

#### After the project's end:

- → Procedures to select data to be preserved
- → File formats for preservation



#### 3.2 What is your data preservation plan?

**Audio recordings of interviews** and **transcripts** will be kept in .wav and .pdf format for a minimum of 10 years on a hard drive in a locked filing cabinet in the PI office.

Consent forms will be kept for 10 years in the same locked filing cabinet.

However, anonymised transcripts of the interviews and focus groups will be distributed under a CC-BY license in .csv format.

Data that cannot be anonymized will not be shared publicly.

# 4. Data Sharing and Reuse

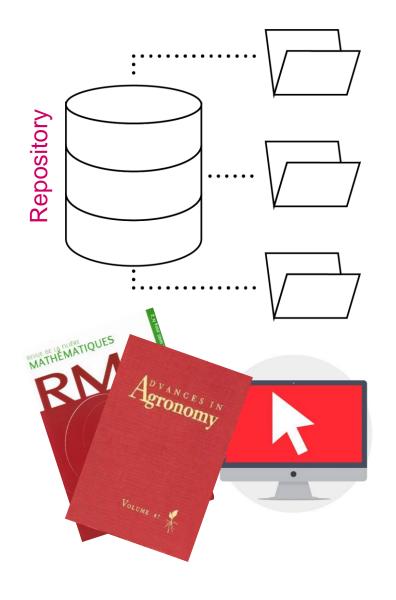
4.1 How and where will the data be shared?

### **Archival and publishing:**

- → Which repository to select
- → How others will find out about your data?



# **Archiving and Dissemination**



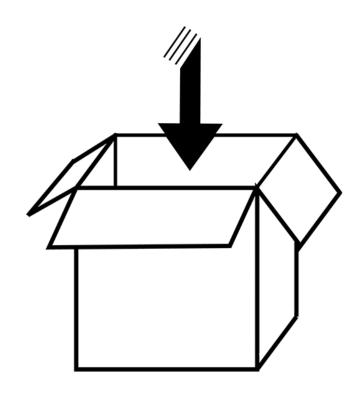
- General purpose (multi-disciplinary)
- ✓ Discipline-specific
- Institutional

- Journal supplementary material service
- Departmental, project or personal web page



# Find a Repository

- Yareta
- Zenodo
- OSF.io
- Dryad
- **...**



For other repositories, check:



### **Yareta**

- Developed as part of a cantonal bill
- Repository for the researchers of Geneva
- ✓ Swiss-based servers (original et copy)



- Compliant with FAIR principles & provides DOI
- Free up to 50 Go (above: 100 CHF / Tb / year of retention)

https://yareta.unige.ch

4.1 How and where will the data be shared?

The project data will be shared with **Yareta**, the research data repository of Geneva's Higher Education Institutions, Along with the README file mentioned in 1.3.

The DOI issued to datasets in the repository will be included as part of a data citation in publications, allowing the datasets underpinning the publication to be easily identified and accessed.

# 4. Data Sharing and Reuse

4.2 Are there any necessary limitations to protect sensitive data?

- → Under which conditions will the data be made available?
- → Timing of data release
- → Delay if applicable

4.2 Are there any necessary limitations to protect sensitive data?

Individual research subjects' data cannot legally nor ethically be made available to non authorised people (HRA, cf. §2.1). Only the sponsor, the investigation team, reviewers, auditors and inspection authorities are entitled to access such data.

No personal data or data that may easily identify subjects will be provided, with respect to the Swiss law on human research (Federal Act on Research involving Human Beings (HRA)) and its applicable ordinance ClinO/KlinV/OClin/OSRUm.

(eg from **DMP Model**, HUG CRC)

# 4. Data Sharing and Reuse

4.3 All digital repositories I will choose are conform to the FAIR Data Principles.

 $\rightarrow \square$  [checkbox]

4.4 I will choose digital repositories maintained by a non-profit organisation.

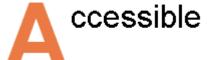
→ yes / no [radio button]

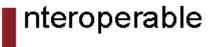


# Is my Repository FAIR?













ark







At least one orange locker

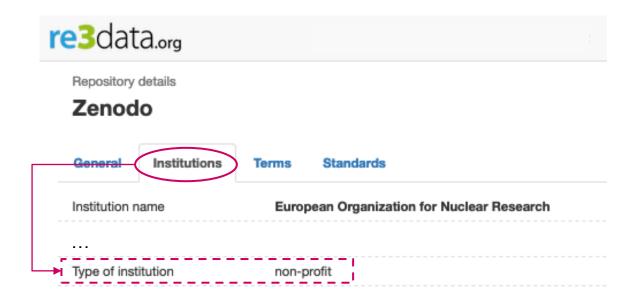
Licence icon | mandatory |

# **Data Management Costs**

The SNF allows you to ask for up to 10'000.- CHF for:

- Data preparation and validation (stewardship)
- Uploading your data on a non-commercial FAIR repository



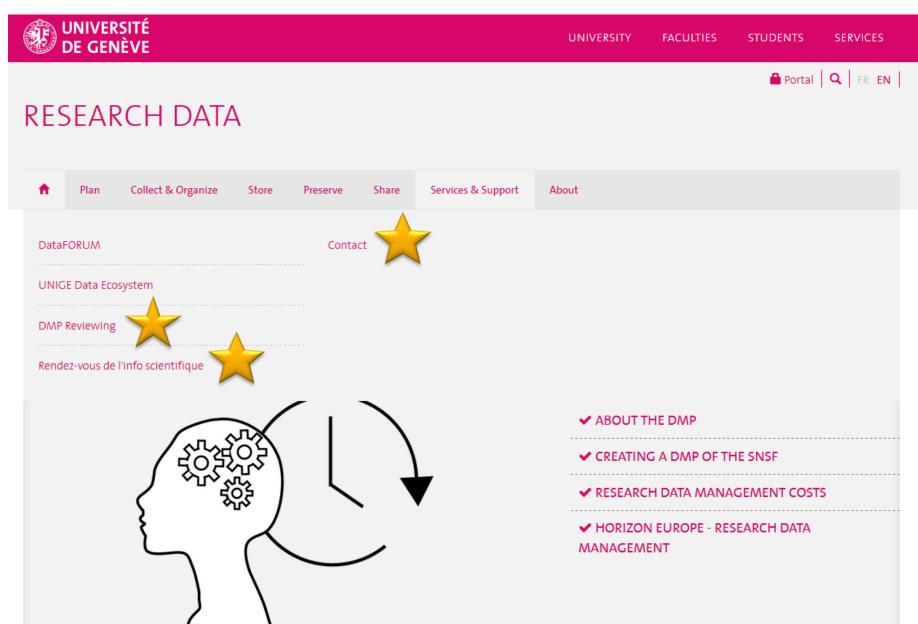




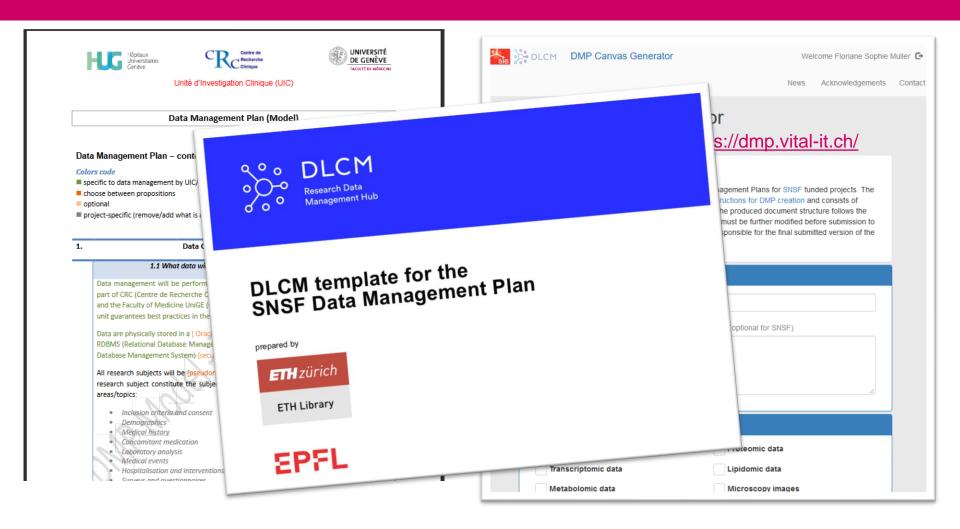
# **Take Home Message**

- ✓ A "plausible" DMP is a sine qua non condition for the release of the funds
- ✓ DMPs are editable throughout the funding period (and must be updated)
- Once SNSF funding has ended and the final scientific report has been approved, the DMP cannot be modified anymore

### www.unige.ch/researchdata



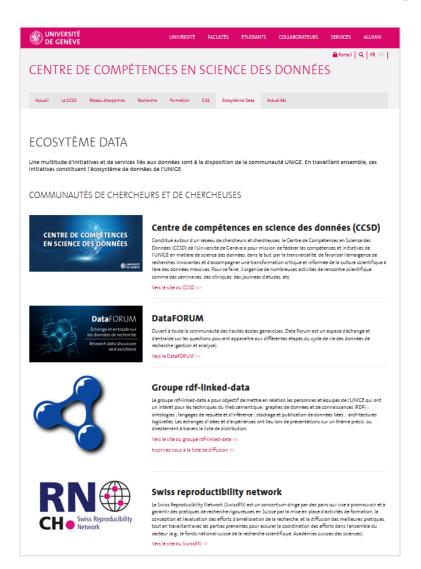
# **DMP Templates**



See: <a href="https://www.unige.ch/researchdata">https://www.unige.ch/researchdata</a> > Plan > Filling a DMP for the SNSF

# **UNIGE's Data Ecosystem**

#### https://datascience.unige.ch/ecosysteme-data-unige



A multitude of data-related initiatives and services are available to the UNIGE community.

Working together, these initiatives constitute the UNIGE data ecosystem

#### **Contacts**

For more informations or help on DMPs and research data:



researchdata-info@unige.ch

http://www.unige.ch/researchdata

Illustrations: CC BY Jørgen Stamp, <a href="http://digitalbevaring.dk">http://digitalbevaring.dk</a>



Bibliothèque de l'UNIGE, 2021

Ce document est sous licence Creative Commons Attribution - Partage dans les mêmes conditions 4.0 International: http://creativecommons.org/licenses/by-sa/4.0/deed.fr.

